

附录 1

定理 1 的证明: 参考张世英和黄违洪 (1987) 的分析方法, 普通最小二乘准则可以用来定义线性模型相应的残差平方和作为分类 $\{u_i, u_{i+1}, \dots, u_j\}$ ($i < j$) 的直径, 即

$$D(i, j) = \sum_{v=i}^j (y_v - \hat{y}_v)^2 = \sum_{v=i}^j (y_v - \mathbf{x}_v^T \hat{\boldsymbol{\beta}})^2 = \sum_{v=i}^j y_v^2 M_l,$$

其中 $M_l = I - \mathbf{x}_v (\mathbf{x}_v^T \mathbf{x}_v)^{-1} \mathbf{x}_v^T$, $\hat{\boldsymbol{\beta}}$ 是对 $\{u_i, u_{i+1}, \dots, u_j\}$ ($i < j$) 的线性模型系数的最小二乘估计值。

将 $n-1$ 个样本分成 k 类, 即 $p(n-1, k): \{i_1, i_1+1, \dots, i_2-1\}, \dots, \{i_k, \dots, n-1\}$ 。定义这一分类的目标函数为 $L[p(n-1, k)] = \sum_{n=1}^k D(j_n, j_n-1) = \sum_{n=1}^k \sum_{v=j_n}^{j_{k+1}-1} y_v^2 M_l (j_{k+1}-1 = n-1)$ 。对该目标函数, 亦有 $L[p(n-1, 2)] = \min_{2, j_n, n-1} \{D(1, j-1) + D(j, n-1)\}$ 和 $L[p(n-1, k)] = \min_{k, j_n, n-1} \{L[p(j-1, k-1)] + D(j, n-1)\}$ 。

接着, 以递推公式计算 $\hat{\boldsymbol{\beta}}$ 和 $\{D(i, j)\}$, 序标从 i 到 j 的线性模型添加了第 $j+1$ 个样本 u_{j+1} , 其中自变量 $\mathbf{x}_{j+1} = (x_{j+1,1}, x_{j+1,2}, \dots, x_{j+1,m})^T$, 则有

$$\hat{\boldsymbol{\beta}}(i, j+1) = \hat{\boldsymbol{\beta}}(i, j) + (\mathbf{y}_{j+1} - \mathbf{x}_{j+1}^T \hat{\boldsymbol{\beta}}(i, j)) \frac{(\mathbf{X}(i, j)^T \mathbf{X}(i, j))^{-1} \mathbf{x}_{j+1}}{\left[1 + \mathbf{x}_{j+1}^T (\mathbf{X}(i, j)^T \mathbf{X}(i, j))^{-1} \mathbf{x}_{j+1}\right]}$$

$$D(i, j+1) = \frac{1}{\left[1 + \mathbf{x}_{j+1}^T (\mathbf{X}(i, j)^T \mathbf{X}(i, j))^{-1} \mathbf{x}_{j+1}\right]}$$

$$(\mathbf{y}_{j+1} - \mathbf{x}_{j+1}^T \hat{\boldsymbol{\beta}}(i, j))^T (\mathbf{y}_{j+1} - \mathbf{x}_{j+1}^T \hat{\boldsymbol{\beta}}(i, j)) + D(i, j)。$$

其中, $\hat{\boldsymbol{\beta}}(i, j) = (\mathbf{X}(i, j)^T \mathbf{X}(i, j))^{-1} \mathbf{X}(i, j)^T \mathbf{y}(i, j)$, $\mathbf{y}(i, j) = (y_i, y_{i+1}, \dots, y_j)^T$,

$$\mathbf{X}(i, j) = \begin{pmatrix} x_{i,1} & x_{i,2} & \cdots & x_{i,m} \\ x_{i+1,1} & x_{i+1,2} & \cdots & x_{i+1,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{j,1} & x_{j,2} & \cdots & x_{j,m} \end{pmatrix}_{(j-i+1) \times m}。$$

当 $j-i, m$ 时, 令 $D(i, j) = 0$ 。

在获得 $\{D(i, j)\}$ 和 $\{L[p(i, j)]\}$ ($1, i, n-1, 1, j, k$) 后, 令 k 为 2, 找到 j 使其满足 $L[p(n-1, 2)] = L[p(1, j-1)] + D(j, n)$, 得到各段 $\{1, 2, \dots, j-1\}, \{j, j+1, \dots, n-1\}$, 则在二分情况下最优分割点 j 存在。

定理 2 的证明: 对 $\tilde{\boldsymbol{\beta}}_{tleve}(\mathbf{w})$ 在 $\mathbf{w}_0 = 1$ 处进行泰勒展开,

$$\tilde{\boldsymbol{\beta}}_{tleve}(\mathbf{w}) = \tilde{\boldsymbol{\beta}}_{tleve}(\mathbf{w}_0) + \frac{\partial \tilde{\boldsymbol{\beta}}_{tleve}(\mathbf{w})}{\partial \mathbf{w}^T} \bigg|_{\mathbf{w}=\mathbf{w}_0} (\mathbf{w} - \mathbf{w}_0) + R_{tleve}$$

其中, 当 $\mathbf{w}$ 逼近 $\mathbf{w}_0$ 时二阶余项 $R_{tleve} = o_p(\|\mathbf{w} - \mathbf{w}_0\|)$ 。设定 $\mathbf{w}_0$ 为单位向量, 即 $\mathbf{w}_0 = 1$, $\tilde{\boldsymbol{\beta}}_{tleve}(\mathbf{w}_0)$ 全样本最小二乘估计 $\hat{\boldsymbol{\beta}}_{ols}$, 即 $\tilde{\boldsymbol{\beta}}_{tleve}(1) = \hat{\boldsymbol{\beta}}_{ols}$ 。则

$$\tilde{\boldsymbol{\beta}}_{tleve}(\mathbf{w}) = \hat{\boldsymbol{\beta}}_{ols} + \frac{\partial (\mathbf{X}^T \text{Diag}\{\mathbf{w}\} \mathbf{X})^{-1} \mathbf{X}^T \text{Diag}\{\mathbf{w}\} \mathbf{y}}{\partial \mathbf{w}^T} \bigg|_{\mathbf{w}=1} (\mathbf{w} - 1) + R_{tleve}$$

通过分部微分, 可得

$$\begin{aligned} \frac{\partial \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{y}}{\partial \mathbf{w}^T} &= \frac{\partial \text{Vec} \left[\left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{y} \right]}{\partial \mathbf{w}^T} \\ &= \left(\mathbf{1} \otimes \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \right) \frac{\partial \text{Vec} \left[\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{y} \right]}{\partial \mathbf{w}^T} \\ &\quad + \left(\mathbf{y}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \otimes \mathbf{I}_p \right) \frac{\partial \text{Vec} \left[\left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \right]}{\partial \mathbf{w}^T} \end{aligned} \quad (1)$$

其中 Vec 是拉直运算, $\otimes$ 为 Kronecker 积, 是任意两个矩阵的直积。

为简化公式 (1) 的第一部分, 可以看作

$$\left(\mathbf{1} \otimes \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \right) \left(\mathbf{y}^T \otimes \mathbf{X}^T \right) \frac{\partial \text{Vec} \left[\text{Diag} \{ \mathbf{w} \} \right]}{\partial \mathbf{w}^T} \quad (2)$$

为简化公式 (1) 的第二部分, 需要以下矩阵微分的结果:

$$\begin{aligned} \frac{\partial \text{Vec} \left[\mathbf{X}^{-1} \right]}{\partial (\text{Vec} \mathbf{X})^T} &= - \left(\mathbf{X}^{-1} \right)^T \otimes \mathbf{X}^{-1}, \\ \frac{\partial \text{Vec} \left[\mathbf{A} \mathbf{W} \mathbf{B} \right]}{\partial \mathbf{w}^T} &= \left(\mathbf{B}^T \otimes \mathbf{A} \right) \frac{\partial \text{Vec} \left[\mathbf{W} \right]}{\partial \mathbf{w}^T} \end{aligned} \quad (3)$$

上述结果的细致过程见 Harville D A (1997) 的第 366~367 页。

整合公式 (3) 的两个结果, 根据链式法则, 可得

$$\begin{aligned} \frac{\partial \text{Vec} \left[\left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \right]}{\partial \mathbf{w}^T} &= \frac{\partial \text{Vec} \left[\left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \right]}{\partial \text{Vec} \left[\left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^T \right]} \frac{\partial \text{Vec} \left[\left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right) \right]}{\partial \mathbf{w}^T} \\ &= - \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \otimes \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \left(\mathbf{X}^T \otimes \mathbf{X}^T \right) \frac{\partial \text{Vec} \left[\text{Diag} \{ \mathbf{w} \} \right]}{\partial \mathbf{w}^T} \end{aligned}$$

通过简单的代数运算, 公式 (1) 推导得到

$$\begin{aligned} &\left\{ \left(\mathbf{y}^T - \mathbf{y}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \mathbf{X}^T \right) \otimes \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \mathbf{X}^T \right\} \frac{\partial \text{Vec} \left[\text{Diag} \{ \mathbf{w} \} \right]}{\partial \mathbf{w}^T} \\ &= \left\{ \left(\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}_{lev}(\mathbf{w}) \right)^T \otimes \left(\mathbf{X}^T \text{Diag} \{ \mathbf{w} \} \mathbf{X} \right)^{-1} \mathbf{X}^T \right\} \frac{\partial \text{Vec} \left[\text{Diag} \{ \mathbf{w} \} \right]}{\partial \mathbf{w}^T} \end{aligned} \quad (4)$$

整合以上结果, 则

$$\begin{aligned}
 \tilde{\beta}_{ilev}(\mathbf{w}) &= \hat{\beta}_{ols} + \frac{\partial \left(\mathbf{X}^T \text{Diag}\{\mathbf{w}\} \mathbf{X} \right)^{-1} \mathbf{X}^T \text{Diag}\{\mathbf{w}\} \mathbf{y}}{\partial \mathbf{w}^T} \Big|_{\mathbf{w}=\mathbf{1}} (\mathbf{w}-\mathbf{1}) + R_{ilev} \\
 &= \hat{\beta}_{ols} + \left\{ \left(\mathbf{y} - \mathbf{X} \hat{\beta}_{ols} \right)^T \otimes \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \right\} \frac{\partial \text{Vec}[\text{Diag}\{\mathbf{w}\}]}{\partial \mathbf{w}^T} (\mathbf{w}-\mathbf{1}) + R_{ilev} \\
 &= \hat{\beta}_{ols} + \left\{ \hat{\mathbf{e}}^T \otimes \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \right\} \begin{pmatrix} \mathbf{e}_1 \mathbf{e}_1^T \\ \mathbf{e}_2 \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_n \mathbf{e}_n^T \end{pmatrix} (\mathbf{w}-\mathbf{1}) + R_{ilev} \\
 &= \hat{\beta}_{ols} + \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \text{Diag}\{\hat{\mathbf{e}}\} (\mathbf{w}-\mathbf{1}) + R_{ilev}
 \end{aligned}$$

其中 $\hat{\mathbf{e}} = \mathbf{Y} - \mathbf{X} \hat{\beta}_{ols}$ 是最小二乘残差向量, $\mathbf{e}_i$ 是一个长度为 n 的向量, 第 i 个元素为 1 其余为 0。

定理 3 的证明: 由于 $E[\mathbf{w}] = \mathbf{1}$,

$$\begin{aligned}
 E_W[\tilde{\beta}_{ilev} | \mathbf{y}] &= E_W[\hat{\beta}_{ols}] + E_W\left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \text{Diag}\{\hat{\mathbf{e}}\} (\mathbf{w}-\mathbf{1})\right] + E_W[R_{ilev}] \\
 &= \hat{\beta}_{ols} + E_W[R_{ilev}]
 \end{aligned}$$

此外, 也容易得到 $E[(w_i - 1)(w_j - 1)] = 1/r\pi_i - 1/r$, $i = j$; $E[(w_i - 1)(w_j - 1)] = -1/r$, $i \neq j$ 。

写成矩阵形式, 则为 $\mathbf{w}$ 的方差 $\text{Var}[\mathbf{w}] = E[(\mathbf{w}-\mathbf{1})(\mathbf{w}-\mathbf{1})^T] = \text{Diag}\{\frac{1}{r\pi}\} - \mathbf{J}_n/r$ 。其中 $\pi = (\pi_1, \pi_2, \dots, \pi_n)^T$, $\mathbf{J}_n$ 是一个 $n \times n$ 单位矩阵。

通过一些额外的代数运算得出 $\tilde{\beta}_{ilev}$ 的条件方差为

$$\begin{aligned}
 \text{Var}_W[\tilde{\beta}_{ilev} | \mathbf{y}] &= \text{Var}\left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \text{Diag}\{\hat{\mathbf{e}}\} (\mathbf{w}-\mathbf{1}) \Big| \mathbf{y}\right] + \text{Var}_W[R_{ilev}] \\
 &= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \text{Diag}\{\hat{\mathbf{e}}\} (\text{Diag}\{1/r\pi\} - \mathbf{J}_n/r) \text{Diag}\{\hat{\mathbf{e}}\} \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} + \text{Var}_W[R_{ilev}] \\
 &= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T [\text{Diag}\{\hat{\mathbf{e}}\} \text{Diag}\{1/r\pi\} \text{Diag}\{\hat{\mathbf{e}}\}] \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} + \text{Var}_W[R_{ilev}] \\
 &= \left(\sum_{j=1}^n h_{jj}^R / r\right) \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \text{Diag}\{\hat{\mathbf{e}}^2 / h_{ii}^R\} \mathbf{X} \left(\mathbf{X}^T \mathbf{X}\right)^{-1} + \text{Var}_W[R_{ilev}]
 \end{aligned}$$

接下来, 继续求解无条件期望与方差。依据迭代期望定律 $E(\mathbf{X}) = E(E(\mathbf{X} | \mathbf{Y}))$, 容易得到 $\tilde{\beta}_{ilev}$ 的无条件均值 $E[\tilde{\beta}_{ilev}] = E[E_W[\tilde{\beta}_{ilev} | \mathbf{y}]] = E[\hat{\beta}_{ols}] + E[E_W[R_{ilev}]] = \beta_0 + E[R_{ilev}]$ 。

最后, 依据方差分解公式 $\text{Var}(\mathbf{X}) = \text{Var}[E(\mathbf{X} | \mathbf{Y})] + E[\text{Var}(\mathbf{X} | \mathbf{Y})]$, 容易得到 $\tilde{\beta}_{ilev}$ 的无条件方差

$$\begin{aligned}
\text{Var}[\tilde{\beta}_{tle}v] &= \text{Var}\left[E_W[\tilde{\beta}_{tle}v|y]\right] + E\left[\text{Var}_W[\tilde{\beta}_{tle}v|y]\right] \\
&= \text{Var}\left[\hat{\beta}_{ols} + E_W[R_{tle}v]\right] \\
&\quad + E\left[\left(\sum_{j=1}^n h_{jj}^R / r\right)(X^T X)^{-1} X^T \text{Diag}\{\hat{e}^2 / h_{ii}^R\} X (X^T X)^{-1} + \text{Var}_W[R_{tle}v]\right] \\
&= \text{Var}[\hat{\beta}_{ols}] + \left(\sum_{j=1}^n h_{jj}^R / r\right)(X^T X)^{-1} X^T \text{Diag}\{\hat{e}^2 / h_{ii}^R\} X (X^T X)^{-1} + \text{Var}[R_{tle}v] \\
&= \sigma^2 (X^T X)^{-1} + \left(\sum_{j=1}^n h_{jj}^R / r\right)(X^T X)^{-1} X^T \text{Diag}\left\{(1 - X\hat{\beta}_{ols} / y)^2 / h_{ii}^R\right\} X (X^T X)^{-1} \\
&\quad + \text{Var}[R_{tle}v] \\
&= \sigma^2 (X^T X)^{-1} + \left(\sum_{j=1}^n h_{jj}^R \sigma^2 / r\right)(X^T X)^{-1} X^T \text{Diag}\left\{(1 - h_{ii}^R)^2 / h_{ii}^R\right\} X (X^T X)^{-1} \\
&\quad + \text{Var}[R_{tle}v]
\end{aligned}$$

推论 1 的证明: Leverage 得分的总和等于协变量维度, 即 $\sum_{j=1}^n h_{jj} = p$ 。由正文公式 (4) 可知, $\sum_{j=1}^n h_{jj}^R = \delta_i x_i (x_i^T x_i)^{-1} x_i^T < x_i (x_i^T x_i)^{-1} x_i^T = \sum_{j=1}^n h_{jj} = p$ 。两阶段 Leverage 重要性抽样的 Leverage 得分总和小于 Leverage 重要性抽样 Leverage 得分总和。

定理 3 中两阶段 Leverage 重要性抽样系数估计的条件方差可以改写为

$$\begin{aligned}
\text{Var}_W[\tilde{\beta}_{tle}v|y] &= \left(\sum_{j=1}^n h_{jj}^R / r\right)(X^T X)^{-1} X^T \text{Diag}\{\hat{e}^2 / h_{ii}^R\} X (X^T X)^{-1} + \text{Var}_W[R_{tle}v] \\
&= \left(\sum_{j=1}^n h_{jj}^R / r\right)(X^T X)^{-1} X^T \text{Diag}\{\hat{e}^2 / \delta_i h_{ii}\} X (X^T X)^{-1} + \text{Var}_W[R_{tle}v]
\end{aligned}$$

在 Ma 等 (2015) 一文中给出了 Leverage 重要性抽样系数估计的方差

$$\text{Var}_W[\tilde{\beta}_{lev}|y] = (p/r)(X^T X)^{-1} X^T \text{Diag}\{\hat{e}^2 / h_{ii}\} X (X^T X)^{-1} + \text{Var}_W[R_{lev}]。$$

因此, 比较以上两式得到 $\text{Var}_W[\tilde{\beta}_{tle}v|y] < \text{Var}_W[\tilde{\beta}_{lev}|y]$ 。

同理, 两阶段 Leverage 重要性抽样系数估计的无条件方差也满足 $\text{Var}[\tilde{\beta}_{tle}v] < \text{Var}[\tilde{\beta}_{lev}]$, 其中 $\text{Var}[\tilde{\beta}_{lev}] = \sigma^2 (X^T X)^{-1} + (p\sigma^2/r)(X^T X)^{-1} X^T \text{Diag}\{(1 - h_{ii})^2 / h_{ii}\} X (X^T X)^{-1} + \text{Var}[R_{lev}]$ 。

参考文献

- [1] 张世英, 黄述洪. 最优分割法的适用性及一类有序样品的聚类方法[J]. 应用数学学报, 1987(2): 138-144.
- [2] Harville D A. Matrix Algebra from A Statistician's perspective[M]. Springer-Verlag, New York, 1997: 366-367.
- [3] Ma P, Mahoney M W, Yu B. A Statistical Perspective on Algorithmic Leveraging[J]. Journal of Machine Learning Research, 2015, 16(27): 861-911.

附表 1

Offset 型异常值数据线性回归结果

方法	正态分布				T_{10} 分布				T_3 分布			
	MSE	R^2	MAE _{coe}	ORA	MSE	R^2	MAE _{coe}	ORA	MSE	R^2	MAE _{coe}	ORA
异常值比例: $\tau = 0.5\%$												
UNIF	1.9234	0.8781	0.0364	0.82	1.8884	0.9046	0.0302	0.82	2.0048	0.9463	0.0206	0.82
LEV	5.5982	0.8227	0.1840	2.58	2.9629	0.9347	0.0765	1.88	1.8519	0.9998	0.0055	1.64
SUBLEV	2.4275	0.9120	0.0603	1.04	1.8175	0.9715	0.0085	0.64	1.8440	0.9996	0.0041	1.20
KLEV	4.9736	0.8268	0.1793	2.76	2.4016	0.9583	0.0364	1.16	2.1609	0.9994	0.0102	0.98
TLEV	1.7230	0.9591	0.0160	0.6	1.6954	0.9834	0.0101	0.46	1.8517	0.9962	0.0047	0.50
异常值比例: $\tau = 2\%$												

UNIF	4.1745	0.7676	0.1377	1.01	4.1819	0.8004	0.1224	1.01	5.1461	0.8677	0.0840	1.01
LEV	12.0432	0.6337	0.4002	2.18	6.3652	0.8519	0.1974	1.85	5.0019	0.9967	0.0175	1.45
SUBLEV	6.5484	0.7684	0.2338	1.33	3.6752	0.9534	0.0345	0.67	5.1427	0.9984	0.0196	1.07
KLEV	6.2781	0.8295	0.1376	1.10	3.8597	0.9678	0.0160	0.15	5.2019	0.9943	0.0154	0.56
TLEV	3.5150	0.9372	0.0619	0.75	3.5448	0.9643	0.0319	0.70	4.4435	0.9914	0.0064	0.97

附表 2

Wave 型异常值数据线性回归结果

方法	正态分布				T_{10} 分布				T_3 分布			
	MSE	R^2	MAE_{coe}	ORA	MSE	R^2	MAE_{coe}	ORA	MSE	R^2	MAE_{coe}	ORA
异常值比例: $\tau = 0.5\%$												
UNIF	1.5468	0.8974	0.0184	0.82	1.6600	0.9141	0.0243	0.82	2.1605	0.9494	0.0228	0.82
LEV	6.0414	0.8440	0.1600	4.06	5.6873	0.8867	0.1350	3.70	5.3200	0.9894	0.0547	3.54
SUBLEV	6.8188	0.8089	0.1699	3.58	7.2049	0.8615	0.1532	3.48	5.3656	0.9902	0.0519	3.38
KLEV	6.5405	0.8271	0.1715	4.2	8.5856	0.8756	0.1839	4.60	4.6646	0.9859	0.0498	3.34
TLEV	1.5066	0.9615	0.0133	1.46	1.5840	0.9775	0.0085	1.92	1.9841	0.9964	0.0056	1.32
异常值比例: $\tau = 2\%$												
UNIF	2.8267	0.8336	0.0651	1.01	3.3575	0.8540	0.0701	1.01	6.4938	0.8868	0.0655	1.01
LEV	16.9251	0.6256	0.4481	3.91	17.9748	0.6776	0.4335	3.94	17.7645	0.9711	0.1292	3.51
SUBLEV	22.4547	0.5594	0.5216	3.50	24.7357	0.6132	0.5074	3.40	14.7153	0.9473	0.1438	3.48
KLEV	19.7637	0.6186	0.5080	4.24	20.2392	0.7174	0.4404	3.75	12.3846	0.9625	0.1305	2.88
TLEV	2.5999	0.9396	0.0339	1.22	2.8957	0.9619	0.0278	1.57	5.1077	0.9873	0.0161	1.91

附表 3

Mixed 型异常值数据线性回归结果

方法	正态分布				T_{10} 分布				T_3 分布			
	MSE	R^2	MAE_{coe}	ORA	MSE	R^2	MAE_{coe}	ORA	MSE	R^2	MAE_{coe}	ORA
异常值比例: $\tau = 0.5\%$												
UNIF	2.3010	0.8742	0.0389	0.82	2.3528	0.9000	0.0350	0.82	2.7651	0.9427	0.0250	0.82
LEV	14.6732	0.7323	0.4106	5.36	12.1506	0.7959	0.3207	5.22	5.2845	0.9954	0.0555	4.84
SUBLEV	10.6669	0.7813	0.3244	4.26	5.7669	0.8871	0.1432	2.80	3.6786	0.9914	0.0240	2.58
KLEV	15.5887	0.7407	0.4389	6.28	7.0463	0.8883	0.1566	3.26	5.5570	0.9809	0.0490	2.86
TLEV	2.3228	0.9496	0.0356	1.44	2.1331	0.9758	0.0244	1.26	2.4792	0.9961	0.0015	1.14
异常值比例: $\tau = 2\%$												
UNIF	5.4169	0.7525	0.1490	1.01	5.8134	0.7793	0.1428	1.01	8.6898	0.8418	0.1114	1.01
LEV	24.3162	0.6706	0.6430	5.02	24.2432	0.6566	0.6063	4.79	18.9461	0.9537	0.1481	4.21
SUBLEV	21.7497	0.6873	0.5967	4.05	14.6227	0.7327	0.3797	3.07	14.9986	0.9666	0.1162	2.76
KLEV	6.7216	0.8632	0.1338	1.40	10.9012	0.8417	0.1969	1.42	18.3518	0.9717	0.1253	2.08
TLEV	4.8844	0.9109	0.1080	1.13	4.8640	0.9368	0.0815	1.41	7.1215	0.9737	0.0342	2.08